

doc ● p#13



ENTERED

1645

RAW SEQUENCE LISTING

DATE: 04/24/2002

PATENT APPLICATION: US/09/196,447B

TIME: 10:56:06

Input Set : A:\Sequence.lst

Output Set: N:\CRF3\04242002\I196447B.raw

SEQUENCE LISTING

RECEIVED

MAY 01 2002

TECH CENTER 1600/2900

4 (1) GENERAL INFORMATION:

6 (i) APPLICANT: Tripp, Cynthia A.

7 Frank, Glenn R.

8 Grieve, Robert B.

10 (ii) TITLE OF INVENTION: NOVEL PARASITIC HELMINTH

11 P22U PROTEINS

13 (iii) NUMBER OF SEQUENCES: 17

15 (iv) CORRESPONDENCE ADDRESS:

16 (A) ADDRESSEE: Sheridan Ross P.C.

17 (B) STREET: 1700 Lincoln St., Suite 3500

18 (C) CITY: Denver

19 (D) STATE: CO

20 (E) COUNTRY: U.S.A.

21 (F) ZIP: 80203

23 (v) COMPUTER READABLE FORM:

24 (A) MEDIUM TYPE: Floppy disk

25 (B) COMPUTER: IBM PC compatible

26 (C) OPERATING SYSTEM: PC-DOS/MS-DOS

27 (D) SOFTWARE: PatentIn Release #1.0, Version #1.30

29 (vi) CURRENT APPLICATION DATA:

C--> 30 (A) APPLICATION NUMBER: US/09/196,447B

C--> 31 (B) FILING DATE: 19-Nov-1998

32 (C) CLASSIFICATION:

34 (viii) ATTORNEY/AGENT INFORMATION:

35 (A) NAME: Connell, Gary J.

36 (B) REGISTRATION NUMBER: 32,020

37 (C) REFERENCE/DOCKET NUMBER: 2618-13-3

39 (ix) TELECOMMUNICATION INFORMATION:

40 (A) TELEPHONE: 303/863-9700

41 (B) TELEFAX: 303/863-0223

44 (2) INFORMATION FOR SEQ ID NO: 1:

46 (i) SEQUENCE CHARACTERISTICS:

47 (A) LENGTH: 913 base pairs

48 (B) TYPE: nucleic acid

49 (C) STRANDEDNESS: single

50 (D) TOPOLOGY: linear

52 (ii) MOLECULE TYPE: cDNA

54 (ix) FEATURE:

55 (A) NAME/KEY: CDS

56 (B) LOCATION: 3..911

59 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

61 GC GAG TTA AAT AGT CGA ATT TCC GGA GTA CAC CGT AAT ACT GCA GGT

47

-RAW SEQUENCE LISTING

DATE: 04/24/2002

PATENT APPLICATION: US/09/196,447B

TIME: 10:56:06

Input Set : A:\Sequence.lst

Output Set: N:\CRF3\04242002\I196447B.raw

```

62   Glu Leu Asn Ser Arg Ile Ser Gly Val His Arg Asn Thr Ala Gly
63       1           5           10           15
65 GCT TTA CAA CGA TTT GCT CTA AAT GGT CAA AAT ACT CTT AAC GAA GGA      95
66 Ala Leu Gln Arg Phe Ala Leu Asn Gly Gln Asn Thr Leu Asn Glu Gly
67       20           25           30
69 TCA AGT TAT GAG CCA AAC GGA CTA TTT GTA TTT TCA GCA ATA AAC GGT      143
70 Ser Ser Tyr Glu Pro Asn Gly Leu Phe Val Phe Ser Ala Ile Asn Gly
71       35           40           45
73 AGC CAT ACT GAT AGC TTA TCT CAG TAT GGT GAA GGA ATA AAT GAA AAT      191
74 Ser His Thr Asp Ser Leu Ser Gln Tyr Gly Glu Gly Ile Asn Glu Asn
75       50           55           60
77 TAT CAT TCT GGA ACT AAT TAT TAT GAT GAA GTA GAA TTA AGA GAT AAA      239
78 Tyr His Ser Gly Thr Asn Tyr Tyr Asp Glu Val Glu Leu Arg Asp Lys
79       65           70           75
81 ACA AAT CAG ACA TCG TAC ATT AAT GGA AAT GAT AAT GGA ATC AAT GGA      287
82 Thr Asn Gln Thr Ser Tyr Ile Asn Gly Asn Asp Asn Gly Ile Asn Gly
83 80           85           90           95
85 AAG GAT GAT GAA GAT CTG GAT GAA TGC TCT GAT CAA GAA TTC CGA TGT      335
86 Lys Asp Asp Glu Asp Leu Asp Glu Cys Ser Asp Gln Glu Phe Arg Cys
87       100          105          110
89 CCA TAT CTA GCT AAA ACA CTT TGT GTT CAT TAT TTG AAA ATA TGC GAT      383
90 Pro Tyr Leu Ala Lys Thr Leu Cys Val His Tyr Leu Lys Ile Cys Asp
91       115          120          125
93 GGT ATT GAT GAT TGT GGT GAT GGA AGT GAT GAA ATG AAC TGT GCT GAT      431
94 Gly Ile Asp Asp Cys Gly Asp Gly Ser Asp Glu Met Asn Cys Ala Asp
95       130          135          140
97 GAT GAA GTG ATA ACA TCA ATA AAT GGT AAC GAA TCA ATC AAT ATC AGA      479
98 Asp Glu Val Ile Thr Ser Ile Asn Gly Asn Glu Ser Ile Asn Ile Arg
99       145          150          155
101 TGT GAT CCG GAT CAA TTT CGA TGT GAA AAT GGA AAA TGT ATC GCA CAA      527
102 Cys Asp Pro Asp Gln Phe Arg Cys Glu Asn Gly Lys Cys Ile Ala Gln
103 160          165          170          175
105 ATT GAT CGA TGT AAT CGA AAA TAT GAT TGT GAT GAT GGT ACA GAT GAA      575
106 Ile Asp Arg Cys Asn Arg Lys Tyr Asp Cys Asp Asp Gly Thr Asp Glu
107       180          185          190
109 ACA ACT TGT GAA TAT TTC GTG CAA GCT TTG CAA CAA GCG AGA GGT GTA      623
110 Thr Thr Cys Glu Tyr Phe Val Gln Ala Leu Gln Gln Ala Arg Gly Val
111       195          200          205
113 ACG GTG CAG GAT AAT GCA ATT CGA GAT GAC GAG ATA CCA AAT TAT ACT      671
114 Thr Val Gln Asp Asn Ala Ile Arg Asp Asp Glu Ile Pro Asn Tyr Thr
115       210          215          220
117 GTA TCC ATG GAA CAG AAA TAC GAT CAA GTA AAG GAA GAT AAG GAG CGG      719
118 Val Ser Met Glu Gln Lys Tyr Asp Gln Val Lys Glu Asp Lys Glu Arg
119       225          230          235
121 CGA ATG CAA GAG GAG GAG GAA CAG GAA AGG CTG AGA GAG TAC GAG GAA      767
122 Arg Met Gln Glu Glu Glu Glu Gln Glu Arg Leu Arg Glu Tyr Glu Glu
123 240          245          250          255
125 CAG ATA CAG GAA AAA TTG AGG CAG GAG GAA GAA AGA GAA CGG CAA GAA      815
126 Gln Ile Gln Glu Lys Leu Arg Gln Glu Glu Glu Arg Glu Arg Gln Glu

```

RAW SEQUENCE LISTING

DATE: 04/24/2002

PATENT APPLICATION: US/09/196,447B

TIME: 10:56:06

Input Set : A:\Sequence.lst

Output Set: N:\CRF3\04242002\I196447B.raw

```

127          260          265          270
129 CAG GAA AGA AGA CAA AAG GAA CGA GAA AGA ATG GAA CAA GAA AGG ATA      863
130 Gln Glu Arg Arg Gln Lys Glu Arg Glu Arg Met Glu Gln Glu Arg Ile
131          275          280          285
133 AGA CAA GAA TAT GAT GAA AAG GAA AGA CAA AGG CAA TAT GCT GAA CAG      911
134 Arg Gln Glu Tyr Asp Glu Lys Glu Arg Gln Arg Gln Tyr Ala Glu Gln
135          290          295          300
137 GC      913
140 (2) INFORMATION FOR SEQ ID NO: 2:
142 (i) SEQUENCE CHARACTERISTICS:
143 (A) LENGTH: 303 amino acids
144 (B) TYPE: amino acid
145 (D) TOPOLOGY: linear
147 (ii) MOLECULE TYPE: protein
149 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:
151 Glu Leu Asn Ser Arg Ile Ser Gly Val His Arg Asn Thr Ala Gly Ala
152 1 5 10 15
154 Leu Gln Arg Phe Ala Leu Asn Gly Gln Asn Thr Leu Asn Glu Gly Ser
155 20 25 30
157 Ser Tyr Glu Pro Asn Gly Leu Phe Val Phe Ser Ala Ile Asn Gly Ser
158 35 40 45
160 His Thr Asp Ser Leu Ser Gln Tyr Gly Glu Gly Ile Asn Glu Asn Tyr
161 50 55 60
163 His Ser Gly Thr Asn Tyr Tyr Asp Glu Val Glu Leu Arg Asp Lys Thr
164 65 70 75 80
166 Asn Gln Thr Ser Tyr Ile Asn Gly Asn Asp Asn Gly Ile Asn Gly Lys
167 85 90 95
169 Asp Asp Glu Asp Leu Asp Glu Cys Ser Asp Gln Glu Phe Arg Cys Pro
170 100 105 110
172 Tyr Leu Ala Lys Thr Leu Cys Val His Tyr Leu Lys Ile Cys Asp Gly
173 115 120 125
175 Ile Asp Asp Cys Gly Asp Gly Ser Asp Glu Met Asn Cys Ala Asp Asp
176 130 135 140
178 Glu Val Ile Thr Ser Ile Asn Gly Asn Glu Ser Ile Asn Ile Arg Cys
179 145 150 155 160
181 Asp Pro Asp Gln Phe Arg Cys Glu Asn Gly Lys Cys Ile Ala Gln Ile
182 165 170 175
184 Asp Arg Cys Asn Arg Lys Tyr Asp Cys Asp Asp Gly Thr Asp Glu Thr
185 180 185 190
187 Thr Cys Glu Tyr Phe Val Gln Ala Leu Gln Gln Ala Arg Gly Val Thr
188 195 200 205
190 Val Gln Asp Asn Ala Ile Arg Asp Asp Glu Ile Pro Asn Tyr Thr Val
191 210 215 220
193 Ser Met Glu Gln Lys Tyr Asp Gln Val Lys Glu Asp Lys Glu Arg Arg
194 225 230 235 240
196 Met Gln Glu Glu Glu Glu Gln Glu Arg Leu Arg Glu Tyr Glu Glu Gln
197 245 250 255
199 Ile Gln Glu Lys Leu Arg Gln Glu Glu Glu Arg Glu Arg Gln Glu Gln
200 260 265 270

```

RAW SEQUENCE LISTING

DATE: 04/24/2002

PATENT APPLICATION: US/09/196,447B

TIME: 10:56:06

Input Set : A:\Sequence.lst

Output Set: N:\CRF3\04242002\I196447B.raw

```

202 Glu Arg Arg Gln Lys Glu Arg Glu Arg Met Glu Gln Glu Arg Ile Arg
203           275           280           285
205 Gln Glu Tyr Asp Glu Lys Glu Arg Gln Arg Gln Tyr Ala Glu Gln
206           290           295           300
209 (2) INFORMATION FOR SEQ ID NO: 3:
211   (i) SEQUENCE CHARACTERISTICS:
212       (A) LENGTH: 1016 base pairs
213       (B) TYPE: nucleic acid
214       (C) STRANDEDNESS: single
215       (D) TOPOLOGY: linear
217   (ii) MOLECULE TYPE: cDNA
219   (ix) FEATURE:
220       (A) NAME/KEY: CDS
221       (B) LOCATION: 3..626
223   (ix) FEATURE:
224       (A) NAME/KEY: 3'UTR
225       (B) LOCATION: 627..1016
228   (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:
230 GT TTT GTT GTA CTA CTC GTT GTT GCA ATA TGG ATT GAA ATG AGC CAA      47
231   Phe Val Val Leu Leu Val Val Ala Ile Trp Ile Glu Met Ser Gln
232     1           5           10           15
234 GGC CAA CAA ATG ATC AAA CAA TGT AAA TGT TCT GAT ATT GCA CCA TGT      95
235 Gly Gln Gln Met Ile Lys Gln Cys Lys Cys Ser Asp Ile Ala Pro Cys
236           20           25           30
238 CAA TTA ACT GCC GTT CAA TCA GTT TTA CCA TGT GCT GAT CAA TGC CAG      143
239 Gln Leu Thr Ala Val Gln Ser Val Leu Pro Cys Ala Asp Gln Cys Gln
240           35           40           45
242 AAA TAT ATT ACT TCA ATT GGT GGT AAT TAT GAT CAA ATT AGT AAC TGT      191
243 Lys Tyr Ile Thr Ser Ile Gly Gly Asn Tyr Asp Gln Ile Ser Asn Cys
244           50           55           60
246 TTT AAA CAG AAA CAA TCA ATT ATA AAT GAT GCT ATG AAA TGT GCT CAA      239
247 Phe Lys Gln Lys Gln Ser Ile Ile Asn Asp Ala Met Lys Cys Ala Gln
248           65           70           75
250 GAT GCT TTC CCA AAT GCA TGC GCA CAA GGT GAA CCA AAA ATG GTA CCA      287
251 Asp Ala Phe Pro Asn Ala Cys Ala Gln Gly Glu Pro Lys Met Val Pro
252   80           85           90           95
254 AAA CGA TTC GGA AAA GGT CTT CAA TTA GCT GTA ATG ACT GAT ATC AAC      335
255 Lys Arg Phe Gly Lys Gly Leu Gln Leu Ala Val Met Thr Asp Ile Asn
256           100          105          110
258 AAA GAA TTA CAA CGA ATG GGA ATA GCA AAT CAA GTT ACT CAA CTA ATC      383
259 Lys Glu Leu Gln Arg Met Gly Ile Ala Asn Gln Val Thr Gln Leu Ile
260           115          120          125
262 TCC CAA GGT CGA CGA TTC TTT AAA TGC TTC CAA TCG TGT ATG ATG AAA      431
263 Ser Gln Gly Arg Arg Phe Phe Lys Cys Phe Gln Ser Cys Met Met Lys
264           130          135          140
266 AAA TTG GGC TCA TGT TCT CCA GAT TGT GGT CTT GAT TTA CCA TCT GAT      479
267 Lys Leu Gly Ser Cys Ser Pro Asp Cys Gly Leu Asp Leu Pro Ser Asp
268           145          150          155
270 AAT GTT ATG GTT CAA ACA GTT AAA AAT TGC GCT CAA AAA AGT GGT ATT      527

```

RAW SEQUENCE LISTING

DATE: 04/24/2002

PATENT APPLICATION: US/09/196,447B

TIME: 10:56:06

Input Set : A:\Sequence.lst

Output Set: N:\CRF3\04242002\I196447B.raw

```

271 Asn Val Met Val Gln Thr Val Lys Asn Cys Ala Gln Lys Ser Gly Ile
272 160                               165                               170                               175
274 CAA ACT GCA TCG GTG CAA GAT CTT TGC TTT TGC GTC GAA CAA GCT GGT      575
275 Gln Thr Ala Ser Val Gln Asp Leu Cys Phe Cys Val Glu Gln Ala Gly
276                               180                               185                               190
278 ATT CGG CAA CTT TCT GAT GTA TGT CCT CGT ATA CAA ATT TTC AAA ACG      623
279 Ile Arg Gln Leu Ser Asp Val Cys Pro Arg Ile Gln Ile Phe Lys Thr
280                               195                               200                               205
282 AAA TGAGTATTGA GAATATTGCA CTAGCAGCAA TCATTATTTT TCTCGAGAAT      676
283 Lys
286 TTTCGCTATC AATAAGTTGG AATATGATTA CAATAATATA TATATTAACT GCAAAAAATCT      736
288 TTCTTCTTCA AAATTATTTT TCATTTTCGCT CTCATAATTG CATGATAATA GTCATAATGA      796
290 AAAACAGGTT TTCTTTTTTT AAAATGATAA CTTCAAACAA ATAGGTATTT CTTGATATAT      856
292 ATATGTATGT ATGTATGTGT GTGTGTGTGT GTGTGTGTAT GTGTGTGTTT GTGTATGTGT      916
294 ATATGTATGT ATGTATGTAT GTATGTATGT ATGTGTAGGA GAAAAGCAAA CTAAACAGTA      976
296 AATGAAAGAA AAAAAATAAGT CAAATAAAAAG TTTGATAATT      1016
299 (2) INFORMATION FOR SEQ ID NO: 4:
301     (i) SEQUENCE CHARACTERISTICS:
302         (A) LENGTH: 208 amino acids
303         (B) TYPE: amino acid
304         (D) TOPOLOGY: linear
306     (ii) MOLECULE TYPE: protein
308     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:
310 Phe Val Val Leu Leu Val Val Ala Ile Trp Ile Glu Met Ser Gln Gly
311 1                               5                               10                               15
313 Gln Gln Met Ile Lys Gln Cys Lys Cys Ser Asp Ile Ala Pro Cys Gln
314                               20                               25                               30
316 Leu Thr Ala Val Gln Ser Val Leu Pro Cys Ala Asp Gln Cys Gln Lys
317                               35                               40                               45
319 Tyr Ile Thr Ser Ile Gly Gly Asn Tyr Asp Gln Ile Ser Asn Cys Phe
320                               50                               55                               60
322 Lys Gln Lys Gln Ser Ile Ile Asn Asp Ala Met Lys Cys Ala Gln Asp
323 65                               70                               75                               80
325 Ala Phe Pro Asn Ala Cys Ala Gln Gly Glu Pro Lys Met Val Pro Lys
326                               85                               90                               95
328 Arg Phe Gly Lys Gly Leu Gln Leu Ala Val Met Thr Asp Ile Asn Lys
329                               100                              105                              110
331 Glu Leu Gln Arg Met Gly Ile Ala Asn Gln Val Thr Gln Leu Ile Ser
332                               115                              120                              125
334 Gln Gly Arg Arg Phe Phe Lys Cys Phe Gln Ser Cys Met Met Lys Lys
335                               130                              135                              140
337 Leu Gly Ser Cys Ser Pro Asp Cys Gly Leu Asp Leu Pro Ser Asp Asn
338 145                              150                              155                              160
340 Val Met Val Gln Thr Val Lys Asn Cys Ala Gln Lys Ser Gly Ile Gln
341                               165                               170                               175
343 Thr Ala Ser Val Gln Asp Leu Cys Phe Cys Val Glu Gln Ala Gly Ile
344                               180                               185                               190
346 Arg Gln Leu Ser Asp Val Cys Pro Arg Ile Gln Ile Phe Lys Thr Lys
347                               195                               200                               205

```

~~V~~ERIFICATION SUMMARY

DATE: 04/24/2002

PATENT APPLICATION: US/09/196,447B

TIME: 10:56:07

Input Set : A:\Sequence.lst

Output Set: N:\CRF3\04242002\I196447B.raw

L:30 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]

L:31 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]